had led to far too much time focused on testing, and another third indicated that this was somewhat the case. Most of the remaining teachers said that the focus is just right. Similarly, two-thirds of the teachers surveyed agreed more with the statement, "State testing is forcing you to concentrate too much on information that will be on the test to the detriment of other important areas" as opposed to "State testing is helping you as a teacher to focus on teaching what children really need to know."

How Much Do Teachers Integrate Standards and Testing Into Their Teaching?

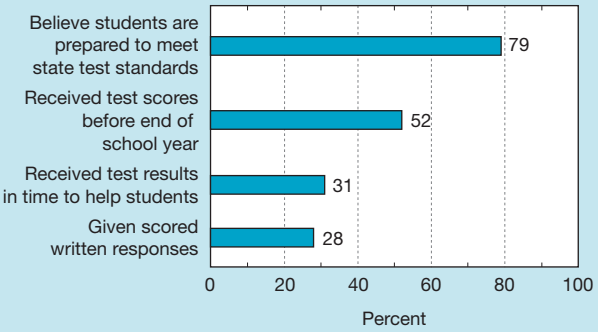
The 2000 *Education Week* survey of public school teachers cited above also indicates that teachers feel prepared to implement state standards in their classrooms, more so than in the previous year (Belden, Russonello, and Stewart Research and Communications 2000). Almost all of the public school teachers (94 percent) reported that they have a copy of the statewide academic standards, and 84 percent said that they have modified their curriculum to reflect the standards (36 percent a "great deal" and 48 percent "somewhat"). A similar proportion said that they have adopted or developed modules, units, or lesson plans linked to the state standards.

A significant amount of "teaching to the test" appears to occur, but using these tests as diagnostic tools is also quite widespread. Nearly 8 out of 10 teachers reported instructing their classes in the previous year in test-taking skills, such as pacing themselves and filling in bubbles clearly (45 percent "a great deal" and 34 percent "somewhat"); 7 out of 10 teachers reported using individual results to help diagnose what students need (36 percent "a great deal" and 34 percent "somewhat"); and 6 out of 10 teachers reported using results to diagnose what they need to be teaching in their classes (32 percent "a great deal" and 42 percent "somewhat"). Nearly two-thirds of teachers said that they had amended what they taught in the previous year to fit what is on the state tests (22 percent "a great deal" and 43 percent "somewhat"). (See sidebar, "High School Teachers Have a Generally Favorable Opinion of State Graduation Tests.") (See figure 1-12.)

While the data in this section have shown that the vast majority of states have adopted content standards in mathematics and science and that state-wide testing in these subjects is increasing, a number of studies raise concerns over the degree to which state tests align with state standards. For example a recent study from the American Federation of Teachers found that "no state or the District of Columbia has a fully developed standards-based system that links quality standards to tests, curriculum and accountability measures" (AFT 2001). This study found that:

- Almost a third of the states' tests are based on weak standards;
- Forty-four percent of those tests are not aligned to the standards;
- Fewer than one-third of the tests are supported by adequate curriculum; and
- One-third of the tests used in decisions regarding promotion or graduation are not aligned to the standards.

Figure 1-12.
Opinion on preparation for and utility of state test by public high school teachers whose state has graduation test: 2000



NOTE: Data are based on responses to the following questions:
Q51. Are students well prepared enough to meet the standards on the tests, or are they ill prepared?
Q52. Last year, did you receive your students' scores on the state exams before the end of the year?
Q53. Last year, did you receive your individual students' test results early enough in the year or too late to be helpful in working with those individuals?
Q55. Are you given copies of your students' scored written responses on the state exams?

SOURCE: Belden, Russonello, and Stewart Research and Communications, *Making the Grade: Teachers' Attitudes Toward Academic Standards and State Testing: Findings of National Survey of Public School Teachers for Education Week* (Washington, DC: 2000).

Science & Engineering Indicators – 2002

While other studies come up with different numbers, the problem of alignment between standards, testing, instruction and accountability remains a common theme (e.g., Achieve, Inc. 2001; CCSSO 2001; Finn and M.J. Petrilli 2000). (See sidebar, "A Survey of Curriculum Use in Classrooms.") Data presented in this section show that both teachers and the general public support standards and testing, although the latter more strongly than the former. The next section examines how the organization of the math and science curriculum in the United States differs from other countries and reviews current measures of the quality of mathematics instruction.

Curriculum and Instruction

Debate continues over the effectiveness of two distinct instructional approaches: (1) emphasis on drill and practice activities in which students work toward skill mastery and (2) emphasis on reasoning, conceptual understanding, and skill application. This debate is driven by differences in opinion regarding the nature of the curriculum as well as different theories about how people learn. Although whole-group instruction and worksheets are still commonly used , the majority of American teachers report using small-group instruction as well as using manipulatives or models to dem-

High School Teachers Have a Generally Favorable Opinion of State Graduation Tests

In the 2000 survey of public school teachers conducted for *Education Week,* a series of questions on testing was asked of public high school teachers who reported that they have a state graduation test. Generally, these high school teachers have favorable opinions of the graduation test.

- A majority (54 percent) believed that the graduation test in their state is appropriate. Only 1 in 10 (13 percent) believed it is too difficult, and 15 percent believed it is too easy. Twenty percent (2 in 10) were unable to offer an opinion of the test.
- A total of 8 in 10 (79 percent) reported that their students are well prepared to meet the standards on the tests. Only 1 in 10 (13 percent) believed that their students are ill prepared.

These high school teachers differed widely, however, on whether the tests are helpful as a diagnostic tool.

- Fifty-eight percent of the teachers reported that test results are helpful for improving their own teaching. Only 1 in 10 (11 percent) found the test results very helpful, and 47 percent said they are somewhat helpful. One-quarter of high school teachers said the results are not at all helpful.

One reason these high school teachers may not find the tests more useful is that the teachers are not receiving the results, or if they are, they are not receiving them in time to implement changes.

- Only half (52 percent) of these high school teachers received their students' scores on the state exams before the end of the year.
- Only 3 in 10 (31 percent) said they received the test results early enough to help individual students.
- Only 3 in 10 (31 percent) were given copies of their students' scored written responses on the state tests.

NOTE: Based on a sample of 173 high school teachers who said their state has a graduation test.

SOURCE: Belden, Russonello, and Stewart Research and Communications 2000.

onstrate a concept (Henke, Chen, and Goldman 1999).[9] Data from the TIMSS video study indicate, however, that teacher implementation of the kinds of instructional techniques for mathematics advocated in the NCTM standards are often superficial. National data that link these approaches to differences in learning outcomes are sparse. This section reviews the most recent data available on curriculum and instruction.

Data from the TIMSS video study show considerable cross-national variation in curricular approaches used in mathematics instruction. For example, American and German middle school mathematics lessons focus primarily on the acquisition and application of skills, but Japanese lessons stress problem solving and thinking. Furthermore, the quality of U.S. mathematics lesson plans was judged to be substantially below that in Germany and Japan in an evaluation by U.S. college mathematics teachers. International studies have also shown that U.S. math and science textbooks cover comparatively more topics with less depth of coverage and development. Recent studies by the American Association for the Advancement of Science (AAAS) have found the most widely used middle school mathematics textbooks and high school science (e.g., biology) textbooks to be less than satisfactory (AAAS 1999a,b and 2000a,b).

Both the new mathematics and the new science standards envision instruction that challenges students, but neither provides an exact blueprint for action. Measuring the extent to which this vision is becoming a reality is difficult because available methods cannot measure quality directly. Instead, educational researchers have relied most often on indicators of the amount of time students spend studying a subject (classwork and homework), the content of lessons, and the types of instructional resources used (e.g., textbooks). This section reviews instructional and curricular topics where recent data collection and research have been strongest: international comparisons of time spent studying mathematics and science, cross-national comparisons of curricular structure, and evaluations of the quality of mathematics and science textbooks. Although these lines of research have yielded valuable information for education policymakers, much remains to be learned about how to make mathematics and science instruction more effective.

Instructional Time

The question of whether U.S. students spend enough time in school or receiving instruction has persisted for many years, and research results on this issue are mixed. Research by Stigler and Stevenson (1991) showed that U.S. students spend fewer hours in school than Japanese students and that U.S. schools allocate less time to core instruction than do other industrialized nations. For example, core academic time in U.S. schools was estimated at 1,460 hours during the four years of high school compared with 3,170 hours in Japan. NECTL reported in 1994 that at the time of the Commission's study, only 10 states specified the number of hours to be spent in academic subjects at various grades. Only eight others provided recommendations regarding academic time. Based on these and other findings, the Commission concluded: "[T]ime is the missing element in the debate about the need for higher academic standards.…We have been asking the impossible of our students—that they learn as much as their foreign peers while spending only half as much time in core academic studies" (NECTL 1994).

[9]Manipulatives are materials designed to provide concrete, hands-on experiences that can help students make the link between math concepts and the real world.

A Survey of Curriculum Use in Classrooms

States' movement toward standards-based reform in mathematics and science has produced strong interest in reliable data for evaluating the effects of reforms. A recent study by the Wisconsin Center for Education Research (WCER) and the Council of Chief State School Officers (CCSSO) applied research-based models and instruments for studying the curriculum to the broader purpose of reporting indicators of curriculum and instruction that could be used by policymakers and educators. States were asked to voluntarily participate in the study if they were interested in gaining information on effects of their reform efforts and gaining knowledge about the development and use of a survey approach to analyzing curriculum. In 1999, schools and teachers in 11 states participated in a study of the enacted curriculum in mathematics and science classrooms. Half the schools selected had high involvement in their state's initiative for improving math or science education ("Initiative" schools), and the other half were schools with less involvement but were similar to the first group based on student demographics ("Comparison" schools). More than 600 teachers across the states completed self-report surveys that covered the subject content they taught and the instructional practices they used in their classes. The enacted curriculum data were designed to give states, districts, and schools an objective method of analyzing current classroom practices in relation to content standards and the goals of systemic initiatives. This National Science Foundation-funded study was a collaborative effort involving state education leaders in science and mathematics, researchers from WCER, and project managers from CCSSO. Educators and researchers worked together to develop survey instruments that would gather reliable data from teachers and students and to develop formats for reporting survey results that would communicate key findings to educators. The goals of the study were to:

- measure differences in instructional practices and curriculum content among teachers and schools,
- determine whether state policy initiatives and state standards lead to differences in math and science teaching, and
- demonstrate the use of "surveys of enacted curriculum" to analyze classroom practices and to produce useful analyses and reports for educators.

The findings from the 1999 study listed below typify the types of issues and questions that can be explored with the survey data.

Active Learning in Science

Question: To what extent are students involved in active, hands-on learning approaches in science class?

- Sample survey data suggest one-fourth of science class time is spent on hands-on science or laboratory activities, but there is wide variation among schools.
- Survey data allow comparison of active science methods in schools that are involved in state initiatives and of science teaching in typical schools.

Problem Solving in Mathematics

Question: To what extent are students in math class learning problem-solving and reasoning skills and learning how to apply knowledge to novel problems?

- A majority of teachers report teaching problem solving in math, but teachers use a wide variety of instructional practices, such as small groups, writing, data analysis, and applying concepts to real-world problems.
- Differences are found in the types and depth of instruction of problem-solving activities between schools involved in state initiatives and comparison schools.

Mathematics and Science Content in Classrooms

Question: How does math and science content taught in classes compare to the goals outlined in state and national standards?

- In middle-grade math and science, most recommended standards are covered, but the level of expectation and depth of coverage vary widely among schools and classes.
- Data reveal differences in the extent of teaching science content across the standards and the extent of articulation between grades.
- Schools differ in their emphasis on algebra, geometry, and data and statistics in the elementary and middle grades.

Multiple Assessment Strategies in Math and Science

Question: What methods of student assessment are used in class, and are the strategies consistent with goals of learning in content standards?

- A majority of teachers use multiple assessment methods in math and science classes but infrequently use extended student responses that require student explanation and justification of answers.
- In science, the survey data allow analysis of differences in the use of performance tasks (hands-on activities) for assessment in class.

Use of Education Technology and Equipment

Question: How is education technology, e.g., calculators and computers, used in math and science instruction? Do teachers have science equipment available in their classes, and how often is it used?

- A majority of elementary- and middle-grade teachers use calculators in teaching math; graphing calculators are available in the typical grade 8 classroom but are rarely used.
- The average elementary school classroom has basic science equipment, but rate of use varies widely among teachers.

Influences on Curriculum and Practices

Question: What effect do state and national standards for science and math learning have on the curriculum taught in classrooms?

- State frameworks and standards and national standards are reported by most teachers to have strong positive influences on their curriculum.
- Survey data allow comparisons of degree of influence on curriculum of state and national standards, textbooks, state and district tests, and teacher preparation and knowledge.

Alignment of Content Taught With State Assessments

Question: Do state assessments reflect what is being taught in classes?

- Analysis of teacher reports and state assessment items shows that tests cover a narrower range of expectations for students than are reported for instruction: tests focus more on memorization, facts, and performing procedures and less on solving novel problems and applying skills and concepts.
- The data on alignment between teacher reports on instruction and content and state assessments allow teachers and assessment staff to examine the areas of weakness and strength of tests and classroom practices.

Teacher Preparation

Question: How well prepared are our teachers to teach science and mathematics?

- The survey data show how well prepared teachers are for using innovative teaching strategies and handling students with varied needs and capacities.
- Middle-grade teachers in math and science receive more professional development than elementary school teachers both in methods of teaching and subject content. Teachers report very positive reactions to professional development related to standards, curriculum, and assessment.

SOURCE: CCSSO 2000b.

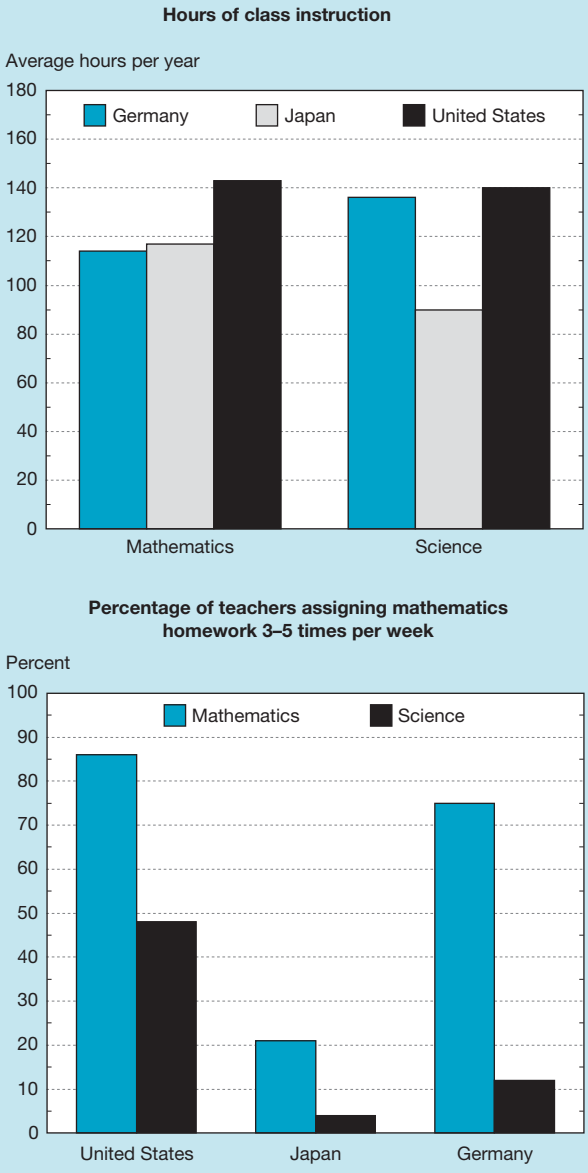
This may not be the case for mathematics and science, as 1995 and 1999 data for 8th graders from TIMSS and TIMSS-R suggest. Eighth-grade students in the United States receive at least as much classroom time in mathematics and science instruction as students in other nations: close to 140 hours per year in mathematics and 140 hours per year in science in 1994-95. (See figure 1-13.) Students in Germany, Japan, and the United States spent about the same amount of time on a typical homework assignment, but U.S. students were assigned homework more often, thus increasing total time spent studying in the two subjects (Beaton et al. 1996b; NCES 1997a,c and 1996c).

Certain caveats are necessary in interpreting results on instructional time. First, in other nations, particularly Japan, students participate in extracurricular mathematics and science activities in afterschool clubs or in formal tutoring activities. Second, disruptions for announcements, special events, and discipline problems in U.S. classrooms considerably reduce the amount of classroom time actually spent on instructional activities (Stigler et al. 1999).

Curriculum and Textbook Content

Analyses conducted in conjunction with TIMSS (Schmidt, McKnight, and Raizen 1997) documented that curriculum guides in the United States include more topics than is the international norm. Most other countries focus on a limited number of topics, and each topic is generally completed before a new one is introduced. In contrast, U.S. curriculums

Figure 1-13.
Selected characteristics of grade 8 mathematics and science instruction, Germany, Japan and United States: 1994–95



NOTE: Data are from the Third International Mathematics and Science Study.

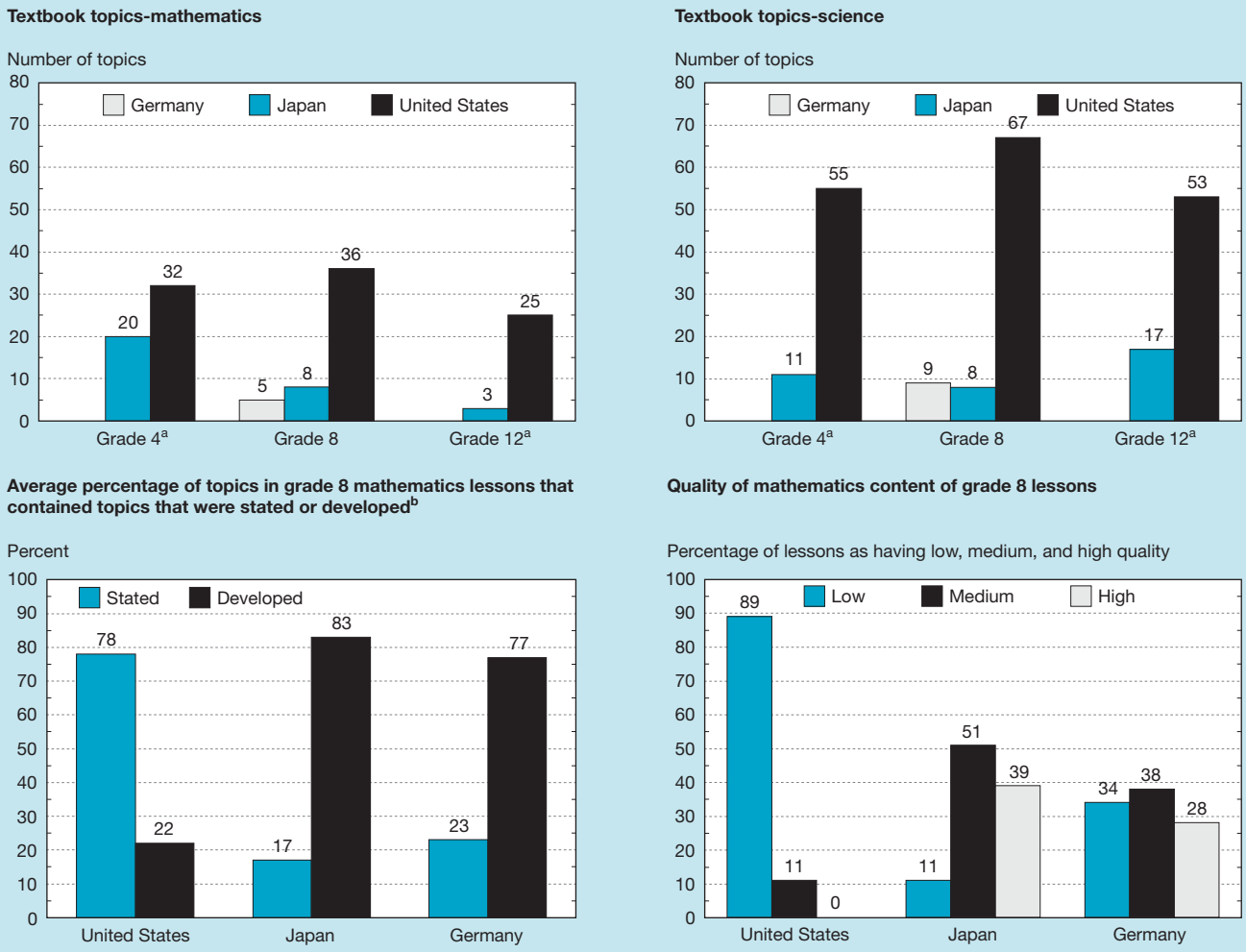
SOURCE: National Center for Education Statistics, *Pursuing Excellence: A Study of U.S. Eighth Grade Mathematics and Science Teaching, Learning, Curriculum, and Achievement in International Context*, NCES 97-198 (Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement: 1996c).

Science & Engineering Indicators – 2002

follow a "spiral" approach: topics are introduced in an elemental form in the early grades, then elaborated and extended in subsequent grades. One result of this is that U.S. curriculums are quite repetitive, because the same topic appears and reappears at several different grades. (See figure 1-14.) Another result is that topics are not presented in any great depth, giving the U.S. curriculum the appearance of being unfocused and shallow.

The Schmidt, McKnight, and Raizen (1997) study also suggests that U.S. curriculums, especially math, make fewer intellectual demands on students, delaying until later grades topics that are covered much earlier in other countries. U.S. mathematics curriculums also were judged to be less advanced, less challenging, and out of step with curriculums in other countries. The middle school curriculum in most TIMSS countries, for example, covers topics in algebra, geometry, physics, and chemistry. Meanwhile, the grade 8 curriculum in U.S. schools is closer to what is taught in grade 7 in other countries and includes a fair amount of arithmetic. Science curriculums, however, are closer to international norms in content and in the sequence of topics. Textbooks reflect the same inadequacies documented by curriculum analyses: insufficient

Figure 1-14.
Selected characteristics of grade 4, 8, and 12 mathematics and science instruction, Germany, Japan, and United States: 1994–95



[a]Data for Germany not available.

[b]A concept was coded as "stated" if it was simply provided by the teacher or students but was not explained or derived. A concept was coded as "developed" when it was derived and/or explained by the teacher or the teacher and students collaboratively in order to increase students' understanding of the concept.

NOTE: Data are from the Third International Mathematics and Science Study. Eighth-grade algebra texts are not included.

SOURCE: J.W. Stigler, P. Gonzales, T. Kanaka, S. Knoll, and A. Serrano, *The TIMSS Videotape Classroom Study: Methods and Findings From an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan, and the United States,* NCES 1999-074 (Washington, DC: U.S. Department of Education, National Center for Education Statistics, Office of Educational Research and Improvement: 1999; W.H. Schmidt, C.C. McKnight, and S.A. Raizen, *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education.* Boston, MA: Kluwer Academic Publishers: 1997).

coverage of many topics and insufficient development of topics. (See figure 1-14.) Compared to textbooks used in other countries, science and mathematics textbooks in the United States convey less challenging expectations, are repetitive, and provide little new information in most grades, a finding reported in earlier research by Flanders (1987) and by Eyelon and Linn (1988). Publishers have made some attempts to reflect the topics and demands conveyed by the educational standards; however, the TIMSS curriculum analyses suggest that when new "standards-referenced" topics are added, much of the old material is retained (Schmidt, McKnight, and Raizen 1997).

Recent studies by AAAS (1999a,b) have reinforced the findings of TIMSS and other studies about the inadequacies of mathematics and science textbooks. AAAS conducted a conceptual analysis of content based on 24 instructional criteria and applied them to the evaluation of 9 middle-school science texts and 13 mathematics texts. The samples included the most widely used texts in both subjects. Each text was evaluated by two independent teams of middle school teachers, curriculum specialists, and science and mathematics education professors. AAAS developed and tested the evaluation procedure over a three-year period in collaboration with more than 100 scientists, mathematicians, educators, and curriculum developers. On a 0- to 3-point scale (where 3 represents "satisfactory"), all nine science textbooks scored below 1.5. Six mathematics texts scored below 1.5, and only three scored above 2.5 points (AAAS 1999a,b).

Similar evaluations of high school biology and algebra texts were only slightly more supportive of their content. In a 2000 evaluation of 10 widely used and newly developed biology textbooks, none received high ratings (AAAS 2000b). Two independent teams of biology teachers, science curriculum specialists, and professors of science education evaluated each biology text, along with its teacher guide. The evaluation examined how well the texts are likely to help students learn the important ideas and skills in the widely accepted Benchmarks for Science Literacy (developed earlier by AAAS Project 2061) and in the National Science Education Standards (NRC 1996). Directors of this study reported, for example, that the textbooks ignore or obscure the most important biological concepts by focusing instead on technical terms and trivial details (which are easy to test) and that activities and questions included are inadequate to help students understand many of the more difficult concepts.

Among the 12 high school algebra textbooks evaluated by AAAS Project 2061, 7 were considered adequate; however, not one was rated highly (AAAS 2000a). Five textbooks, including three that are widely used in American classrooms, were rated so inadequate that they lack potential for student learning. Highlights of the evaluation included the following:

- All of the textbooks present algebra using a variety of contexts and give students appropriate firsthand experiences with the concepts and skills.
- Most of the textbooks do an acceptable job of developing student ideas about algebra by representing ideas, demonstrating content, and providing appropriate practice.
- No textbook does a satisfactory job of providing assessments to help teachers make instructional decisions based specifically on what their students have or have not learned.
- No textbook does a satisfactory job of building on students' existing ideas about algebra or helping them overcome their misconceptions or missing prerequisite knowledge.

Instructional Practice

Most information about instructional practice has come from surveys that asked teachers about specific aspects of their teaching. In a recent survey, 82 percent of full-time U.S. mathematics teachers and 74 percent of full-time science teachers gave themselves good grades on using practices consistent with educational standards in their fields (NCES 1999d). However, classroom observational studies, which have provided more depth and dimension to depictions of practice, often paint quite a different picture. These studies demonstrate that it is relatively easy for teachers to adopt the surface characteristics of standards-based teaching but much harder to implement the core features in everyday classroom practice (Spillane and Zeuli 1999; Stigler et al. 1999; and NCES 2000d).

The TIMSS video study of 8th-grade mathematics instruction is a case in point. Lessons in U.S., German, and Japanese classrooms were fully documented, including descriptions of the teachers' actions, students' actions, amount of time spent on each activity, content presented, and intellectual level of the tasks that students were given in the lesson (Stigler et al. 1999). These findings identified four key points:

- The content of U.S. mathematics classes requires less high-level thought than classes in Germany and Japan.
- The typical goal of U.S. mathematics teachers is to teach students how to do something, but the typical goal of Japanese teachers is to help them understand mathematical concepts.
- Japanese classes share many features called for by U.S. mathematics reforms, but U.S. classes are less likely to exhibit these features.
- Although most U.S. mathematics teachers report familiarity with reform recommendations, relatively few apply the key points in their classrooms.

Ratings by mathematicians of the quality of instruction in 8th-grade German, Japanese, and U.S. mathematics classrooms in 1994–95 suggest a lower level of quality in U.S. instruction. Approximately 30 percent of lessons in Japanese classrooms were rated as "high quality" and 13 percent were rated as "low quality." In German classrooms, 23 percent of lessons received high ratings and 40 percent received low ratings. In comparison, approximately 87 percent of U.S. lessons were considered "low quality" and none were considered "high quality." (See figure 1-14.) However, because of the small scale of the study, these results are suggestive rather than definitive. The studies are now being replicated on a larger scale in both mathematics and science.